Fluent

How Hard Can It Be?

The reality of building a text-to-SQL solution in-house

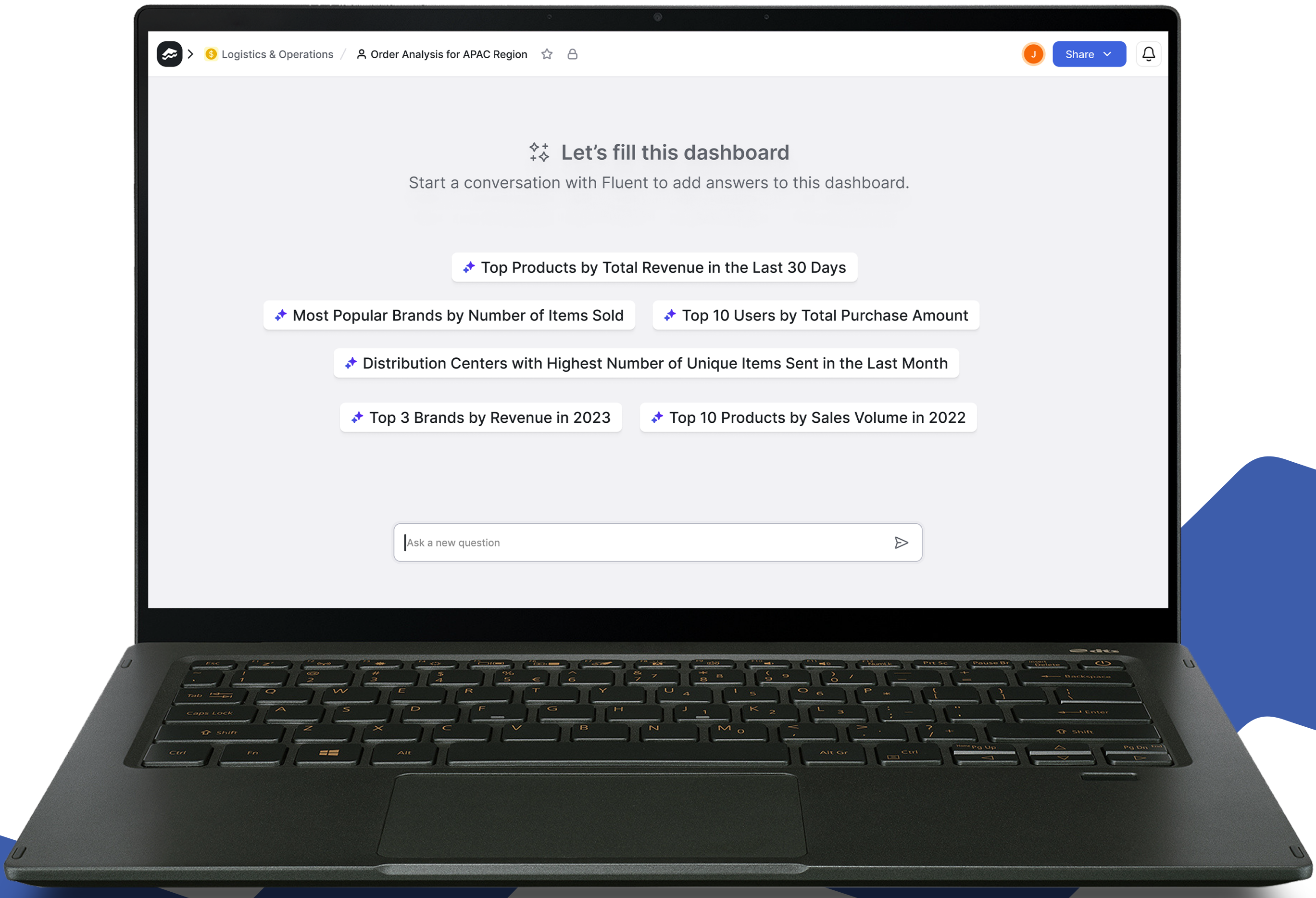


TABLE OF CONTENTS

Executive **Summary**

Describing AI as a 'paradigm shift' and 'game changer' seems a bit tired now doesn't it? We've all agreed that AI is good for business, whether it's designing mattresses, toothbrushes, perfume or fanfiction, or any number of everyday efficiencies.

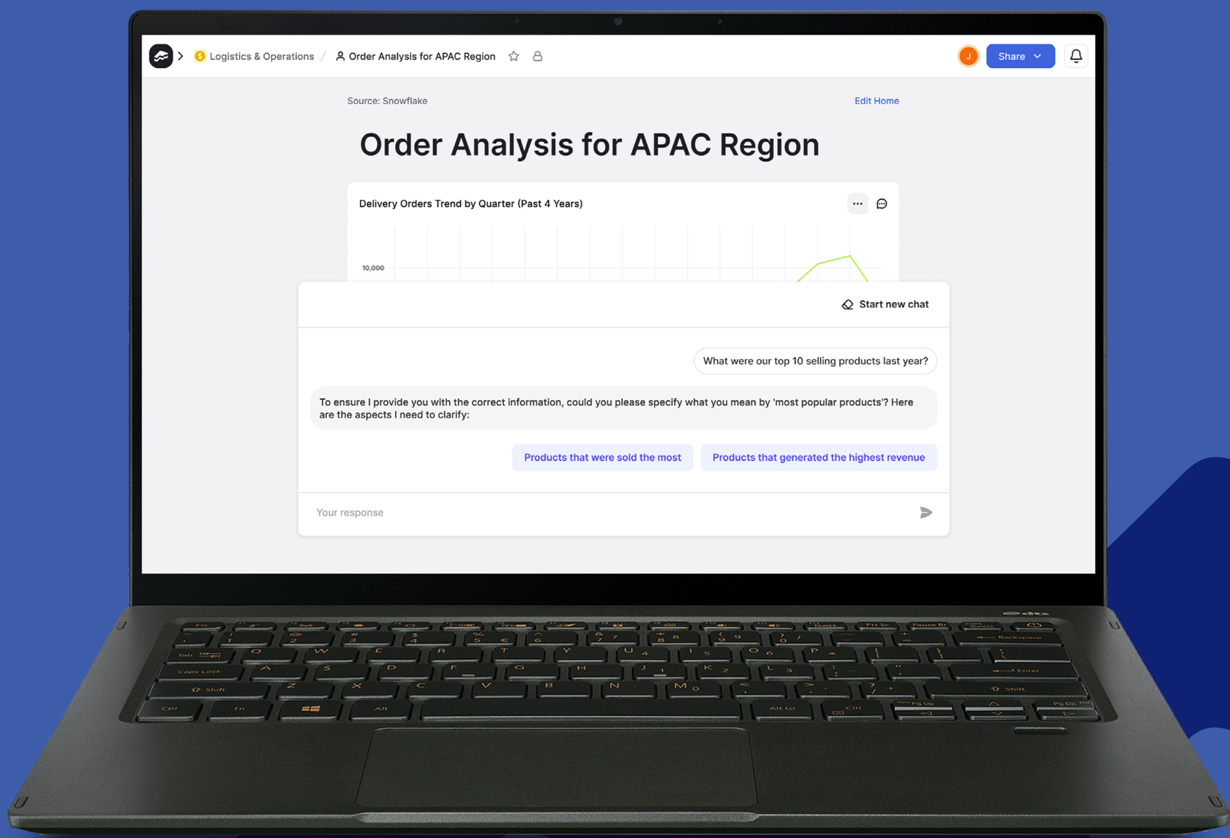
AI solutions continue to evolve and captivate the corporate and consumer world. The number of businesses embracing AI has soared in 2024, following half a decade of flattening levels of adoption. It goes further, too. Half of businesses already deploying AI, are doing so in two or more of their internal functions. Toothbrush designs aside, of AI's many business applications, the use of Large Language Models (LLMs) and SQL generation in data management is making waves. As of late 2023, 58% of businesses were experimenting with LLMs, though only 23% had chosen to 'commercialise' them, according to a recent survey. It's a growing practice, but there's seriously untapped potential.

Enterprises, in particular, sit on mountains of structured commercial and performance data. In simple terms, their data warehouses can be unlocked with the use of LLMs, and translated into plain English with SQL generation. Business users can converse with, and self-serve insights from, the vast data warehouse the enterprise sits on. The internal impact can be transformative. However, how enterprises reach this AI-powered paradise is another question.

Enterprises will spend big on their data. Anything from complex BI tools, data teams, infrastructure, maintenance and implementations. Here's your hard-hitting stat: a projected $59.7 billion will be spent on business intelligence tools in 2024. After sinking millions into these areas, enterprise data teams often end up being inundated with ad-hoc queries from other teams anyway. AI solutions like text-to-SQL agents have the ability to quickly filter and translate vast data warehouses into bite-sized conversational snippets makes it an understandably attractive alternative to the chaotic conveyor belt of queries. Data analysts serve non-technical users with dashboards based on what is being asked directly, or what they think might be needed in future. Lead times for ad-hoc data requests usually number in the days. It's an inefficient system, but it's also a waste of a talented analyst team's time. Implementing LLMs is usually done one of two ways; acquiring a solution externally, or building a SQL solution in-house, with OpenAI, Anthropic or Google's Gemini. Many enterprises opt for the latter.

And who can blame them? In-house IT builds may have historically offered some benefits. There's an assumption it will mean greater cost control and solution tailoring. Remember, this isn't standard software engineering. Evidence suggests that these in-house AI projects hardly ever run smoothly, and carry more complexity, unpredictable cost and scope creep than is ever predicted beforehand.

In this whitepaper, we start by looking at the underlying numbers and perceptions of AI usage in data management right now. We'll walk through theoretically building a natural language query solution (and just how hard it is) and finally, the emerging, alternative means of accessing these capabilities.

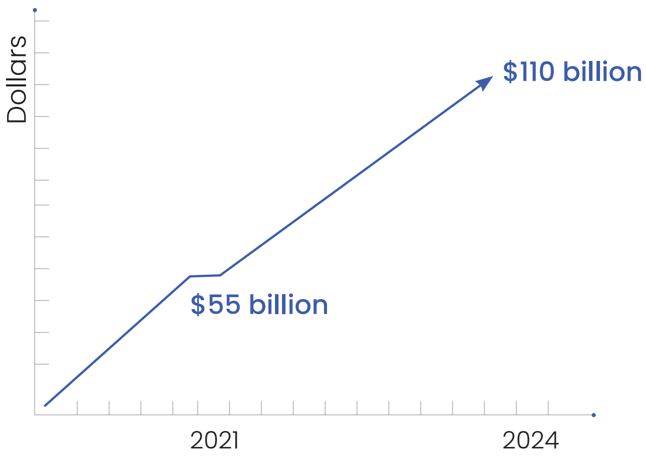


SECTION I

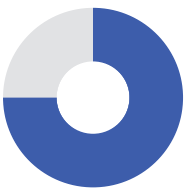
The State of Play

The numbers tell a clear story; most enterprise-level executives see AI as a part of their strategy and tech stack going forward.

The IDC (International Data Corporation) forecasts global spending on AI systems will have doubled since 2021, reaching $110 billion this year - much of it on in-house projects.



With the democratisation of AI, spearheaded by organisations like OpenAI and ChatGPT, it's never been easier to integrate an LLM or Generative AI tool into an enterprise. When Forbes asked 600 US-based business owners for their thoughts on ChatGPT, 97% felt it would help their business in 'some way'. That statement speaks volumes about the perception of AI amongst business leaders. ChatGPT offers many time-saving benefits to customer communication and content creation. There's evidence that even the mere mention of 'AI' in marketing copy increases a customer's willingness to pay by 17%. Sortlist did a study of tech 'giants' who had introduced an AI integration as part of their customer-facing products. Adopters' share price outperformed their competitors by 3.5% on the Nasdaq 100 Tech Sector.



75%

Gartner predicts that by 2025, 75% of enterprises will have shifted from piloting to operationalising AI.

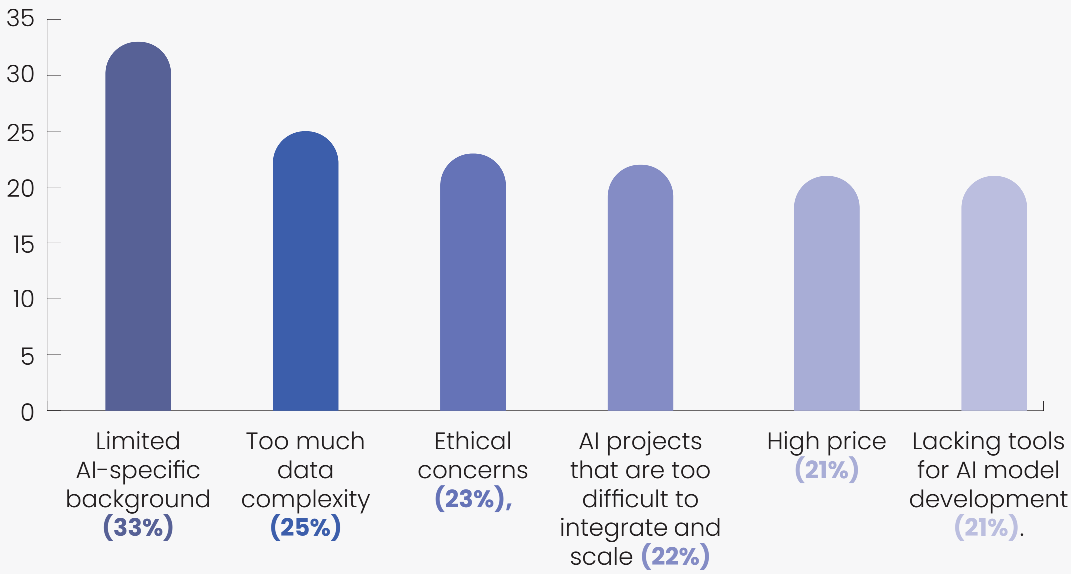
72%

72% of business executives PWC spoke to believe AI will be the 'business advantage of the future.'

Despite all the positive perceptions, and the broad appeal of building AI projects internally, here's the catch.

The vast majority of in-house projects fail.

In 2023, Harvard Business Review estimated that an eye-watering [80% of corporate AI projects fail per year]. That's almost double the rate of corporate IT project failures recorde d a decade prior.



Many C-level executives are overwhelmingly positive about AI in general, but they're struggling to put it into practice, often due to the complexity of data management and analysis. Enterprise-level text-to-SQL solutions need to accurately give non-technical users digestible, explainable data insights. And it's often the sophisticated process of providing explainability that causes problems during in-house builds.

> I think our non-technical users would have even been okay with 85% accuracy. That's not the problem; they needed to understand if something went wrong. With non-technical users, the ability and explainability becomes a lot more important because they can't read the SQL, but they have to be able to figure out those 10 to 15% inaccurate results. They have to be able to catch those things.
>
> **Real customer feedback from a business reflecting on a failed in-house build**

Clarity and explainability when generating answers in plain english is key to a usable solution. An SQL generation tool that needs significant oversight from the governing data team isn't solving the 'ad-hoc query' problem. Just a few inaccurate responses or failure to explain a generated answer creates a lack of trust in non-technical users and we know where that leads: continued reliance on data teams. In some cases, the projects are simply abandoned as a result. But let's say you were to try it. Next up, we'll explore the theoretical build of an enterprise-level text-to-SQL solution, using an LLM.

SECTION II

How To Build An AI Solution In-House - In Theory

Legacy systems, business needs, data quality and resourcing can all impact project length. It will take significant resourcing and expertise to build and integrate an LLM, and launch a text-to-SQL solution, within an enterprise's data function from scratch. In-house projects like this will take 12-20 months, as an optimistic estimate.

A rough roadmap to adoption

01

Initial Planning and Requirement Gathering (1-2 months)

To start, you'll need to scope out the fundamental requirements of your project, what 'good' looks like and project-wide details like budget, resourcing and technical demands.

02

Data Preparation and Infrastructure Setup (2-3 months)

Once initial planning is complete, getting your data in order, from collating to cleaning, ensures you can build on consistent, 'structured' data. This process will impact the effectiveness of your text-to-SQL in the long run.

Considerations

- **Data Collation**
- **Data Cleaning and Transformation**
- **Data Integration** (consider centralising into a data warehouse),
- **Infrastructure Setup** (including databases, servers, and cloud services)

03

Design and development of natural language processing engine (3-6 months)

The longest and most complex phase of the project is building the Natural Language Processing engine. It needs to recognise queries, intents, context and provide consistent and accurate answers. Where you may have found the average internal software project to be unpredictable and labour-intensive, building an NLP engine will be worse, due to the broad complexity featured.

Considerations

- **Natural Language Understanding (NLU), Intent Recognition** (building models to recognize user intents and map them to corresponding BI actions)
- **Entity Recognition**
- **Machine Learning Models Training**
- **Algorithm Development** (to translate natural language queries into structured database queries (e.g., SQL).

04

Development of User Interface (2-4 months)

05

Testing and validation (2-3 months)

As you're working with natural language processing, testing will be significantly more unpredictable, unreliable and complex.

06

Deployment and Business user upskilling (1-2 months)

07

Maintenance and continuous improvement (Ongoing)

As you've opted to build in-house, continued resourcing will need to be allocated to upkeep and maintenance of the solution.

Unlike some Machine Learning Models, LLMs specialise in natural language processing. It makes them a sophisticated variant of an MLM. Building a SQL generation tool using an open source LLM includes a need for good-quality, structured, underlying data and lots of questions around integration. It's these things that can lead to delays and challenges for an in-house project team.

In practice, this rough in-house project roadmap is likely to include a lot more nuance, time and expense. All of that uncertainty is why enterprises can and will seek out alternatives options, too.

SECTION III

Emerging Alternatives

The 2024 McKinsey 'State of AI' study showed that AI inaccuracy, cybersecurity risk and answer explainability have remained key concerns for enterprises year-on-year. We've seen the adoption of AI keep ticking up, but it's matched by growing understanding of the risks and limitations involved, particularly with Generative AI. A combination of caution and readily available open source tools are likely a big driver of in-house builds, over third-party providers.

However, third-party providers know this. Their solutions are increasingly targeted towards servicing enterprise-level clients through a number of features.

How third-party providers are matching enterprise expectations

Offering easily governable LLMs	Offering case-specific customisation features	Establishing enterprise-grade security measures; ie, SOC2 Certification.	Proven scalability and a track record of working with larger clients

Providers with these qualities are able to put the typical enterprise concerns at ease, while also being able to integrate and scale solutions much faster. A market-ready third party solution may take just weeks to become operational. In short, in-house building means being accountable for every technical glitch, delay and cost. Using a third-party can still offer the same controls and security an enterprise data team wants, but with significantly less of the complexity in adopting an LLM, and building on it.

The benefits of third-party providers often centre on cost and time savings. This can absolutely apply to an enterprise. However, the competitive advantage largely lies in how a third party provider can handle the majority of the integration and testing themselves. It means data teams in-house can direct their efforts on longer-term, higher-impact work, and pass over high volume ad-hoc querying to the SQL generation solution they've chosen to partner with. In the long run, they get the same result of a powerful data tool that makes quick, unplanned insight generation effortless, but don't have to make the LLM project their team's core functionality. It becomes a sophisticated solution that quickly deploys and lightens the ad-hoc query load almost entirely.

The Future

We've come a long way from the initial buzz that surrounded the first AI chatbots. The next evolution for data management may well be an increased openness to seeking third-party support in LLM adoption. Given the choice between year-long projects filled with delays and testing, or fast secure integration with a third-party, we predict that the latter will make more business sense moving forward.
Based on the findings in this report, we can produce a few key takeaways for readers working in enterprise-level businesses dealing with some of the challenges we've outlined:

- LLMs and Text-to-SQL solutions are the most efficient way to tackle ad-hoc queries that overload data teams.
- Whether an enterprise chooses to build in-house or not, embracing AI within their data-managing teams is likely to become a critical competitive edge.
- Businesses embracing AI already are beginning to see the benefits and ROI - this trend is likely to continue.
- In-house builds are offering less benefit to enterprises, in light of mature third-party providers who can provide tried and tested text-to-SQL solutions faster, with enterprise-grade scalability and security requirements.

There's always going to be a side of AI that's inconceivably big; the idea of 'endless possibility'. In contrast, the application of AI within data management isn't unchartered territory. Many businesses, including those Fluent work with already, are reaping the benefits.

Generative AI will continue to solve problems for businesses and consumers, some simpler than others. It's just complex to build a conversational layer, powered by an LLM, over a legacy data warehouse in a large business. Thanks to the rise of established third-party providers - who truly understand the unique needs of enterprises - available today, it doesn't have to be.

Contact **us**

If any of the above sounded familiar, or you'd like to learn more - reach out for a conversation with one of our team.

Rob Van Den Bergh
CEO & Co-Founder
Fluent
rob@fluenthq.com

Will Bloomfield
SDR Lead
Fluent
will@fluenthq.com

Endnotes

https://newsroom.ibm.com/2024-02-26-New-IBM-Study-Data-Reveals-74-of-Energy-Utility-Companies-Surveyed-Embracing-AI

https://hbr.org/2023/11/keep-your-ai-projects-on-track

https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-generative-artificial-intelligence.pdf

https://www.forbes.com/advisor/business/software/ai-in-business/#how_businesses_are_using_artificial_intelligence_section

https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

https://springsapps.com/knowledge/large-language-model-statistics-and-numbers-2024#:~:text=In%20a%20survey%20conducted%20in,or%20have%20already%20done%20so